JOHNS HOPKINS
U N I V E R S I T Y

Center for Clinical Trials

Department of Biostatistics
Department of Epidemiology
Department of International Health

Department of Medicine
Department of Ophthalmology
Oncology Center

(Tuesday 8:54am) 10 October 2000

Memorandum

To: Center for Clinical Trials faculty and staff

Fr: Curt Meinert

Re: Stratification good practice policies and procedures (GPPP)

Definitions

stratification *n* - 1. Broadly, the act or process of **stratifying** (defn 2 or 3). 2. An active ongoing process of **stratifying**, as in the sense of defn 2, as in placing **patients** into **strata** as they arrive at a **clinic** as a prelude to **enrollment** and **randomization** to **treatment** in a **trial**. 3. The act or process of classifying **treatment units** or **observations** into strata after **enrollment** for a **subgroup analysis**; **post-stratification**. rt: **classification**, **quotification** ***Usage note***: Stratification is done as a means of **controlling** sources of **variation** related to or assumed to be related to the **outcome**. Stratification (defn 2) and **blocking** in the **treatment assignment** process serve different purposes. Blocking is imposed as a means of ensuring that the **assignment ratio** will be satisfied or nearly satisfied; stratification is done to ensure the comparability of the treatment groups with regard to the variable(s) used in stratification. There is confusion regarding the meaning and impact of stratification on the design and operation of a trial. Often the act of stratification is taken as evidence of the need to perform **treatment comparisons** within the various strata represented in the stratification. Although that may be desirable, such comparisons are not necessary. Valid comparisons of the **treatment groups** can be performed by pooling across strata. As a rule, the mix of persons enrolled into a trial is determined by the mix of persons seen and ultimately judged eligible for **enrollment**. Hence, the numbers to be represented in the various strata will be **variables** having values known only after completion of enrollment. The imposition of a **sample size requirement** for one or more of the strata (see **recruitment quota**), in addition to one for the trial, extends the time required for **recruitment** and should not be imposed unless there are valid scientific or practical reasons for doing so. Confusion also arises from use of the term **stratification** in two distinctly different contexts, as suggested in defn 2 and 3 above. Use **post-stratification** for uses in the sense of defn 3, especially when in settings, such as trials, where both forms of stratification are used.

stratification variable *n* - 1. A **variable** used to classify **treatment units** into **strata** in relation to **treatment assignment**. 2. A variable used to classify **observation units** into strata in relation to **data analysis**.

P&P 1: Stratify sparingly.

Room 5010 • 615 North Wolfe Street • Baltimore, Maryland 21205 • 410 955-8175 • Fax: 410 955-0932
E-mail: cctmail@jhsph.edu • Web site: http://www.jhsph.edu/Research/Centers/CCT

Comment

Stratification is overrated as a means of variance control. At best, it can serve only to control a few selected variables. One still has to trust to "luck" for unbiased distribution of other baseline and demographic variables across treatment groups. Hence, one still needs to explore results for baseline imbalances across treatment groups.

The statistical gain, via reduced variance, is small in trials involving sample sizes of 100 or more per treatment group.

P&P 2: Avoid "overstratification".

Comment

The utility of stratification is a function of the extent to which it ensures baseline comparability across treatment groups for the variable or variables selected. The more strata the greater the likelihood of imbalances remaining even with stratification.

Operationally, the number of strata should be small relative to the finished sample size.

P&P 3: Avoid stratification using variables not directly observed.

Comment

Variables not directly observed can delay randomization because assignments cannot be obtained until the information needed for stratification is in hand. Such delays (eg, the delay in waiting for results of a laboratory test needed for stratification) are avoided when the choice of variables is limited to the set amenable to direct observation, that is, variables that can be measured or assessed by observation, examination, or report of the patient (eg, gender, age, ethnic origin, body weight, or disease history).

P&P 4: Steer clear of stratification variables that cannot be measured or ascertained well within the time frame needed for use in stratification.

Comment

Obviously, if a variable is to be used for stratification its value has to known by the time of randomization. Therefore, it is usually a good idea to steer clear of laboratory tests for stratification, unless they can be done a day or two of when a person was seen.

If pathology readings are used for stratification, the stratification is best done using local readings. Use of procedures involving two or more independent readings and adjudication by a third party in the case of disagreements is usually best avoided because of the likelihood of delaying randomization.

P&P 5: Limit stratification so as to be unlikely to have strata, when enrollment is finished, of size less than 2 times (or some larger multiplier) of the minimum block size (eg, 15 in a trial like the CDP having 5 test treatments and a control treatment and an assignment ratio of 1:1:1:1:1:1.5).

Comment

The purpose is to force planners to be parsimonious in regard to stratification and to avoid "empty" or nearly empty strata.

P&P 6: Proposals for use of a stratification variable as a vehicle for variance control should not be accepted in the absence of evidence suggesting that the state of the variable influences or moderates the treatment effect or is related to outcome.

Comment

The purpose is the same as for P&P 5.

P&P 7: Any variable producing or suspected of producing quantitative treatment interactions should be controlled by stratification.

Comment

Note that the reference is to quantitative interaction. The proper strategy in the case of known or suspected qualitative interactions is to exclude the subgroup of persons not likely to benefit from treatment.

P&P 8: Avoid stratification schemes subject to classification error.

Comment

The purpose of stratification is for variance control. The utility of stratification is reduced when persons cannot be accurately classified as to strata. A high error rate in classification, even if not of serious consequence in regard to the randomization process, will reflect badly on the trial when revealed.

P&P 9: Avoid confusing stratification with quotification.

Comment

The tendency for the uninitiated is to assume that stratification implies a recruitment strategy aimed at enrolling a specified number of persons per stratum. It does not. Its only purpose is to ensure, within the limits of the stratification, the same mix of people with respect to the stratification variable(s) across treatment groups. It does not ensure a specified number of persons by stratum.

P&P 10: Avoid confusing stratification with mandative subgroup analysis.

Comment

The uninitiated tend to assume that stratification obligates one to estimate treatment effect by stratum. It does not, though, often, results at least for the primary outcome are presented overall and by stratum.

P&P 11: Variables used to restrict or adjust treatment should be used for stratification. For example, gender, in the case of a trial where the menu of treatments or dosage level is different for males vs females, should be a stratification variable.

P&P 12: Generally, clinic in multicenter trials is used for stratification.

Comment

Clinic populations can be strikingly different. Stratification by clinic avoids confounding of treatment effect by clinic. Clinic also has logistical value as a stratification variable in drug trials. Randomization by clinic enables the coordinating center to forecast drug needs by clinic.

P&P 13: Avoid choosing variables merely on the supposition of being related to treatment.

Comment

The purpose is as stated for P&P 5. Investigators have no trouble speculating as to ways in which variables influence treatment or outcome. The list of candidate variables will be unrealistically long, absent requirements of evidence supporting the speculations.